THE ATLAS PIXEL BARE MODULE

Abstract

This is a comprehensive document describing the ATLAS Pixel bare module. It is written as documentation for the Production Readiness Review and for internal future reference.

Prepared by:	***Checked by:***	***Approved by:***
A. Andreazza, K. Einsweiler, M. Garcia-Sciveres, C. Gemme, D. Giugni, F. Huegging, P. Netchaeva, M. Olcese, J. Richardson, L. Rossi, A. Rozanov, C. Troncon, T. Stockmans, N. Wermes		

Distribution List

Table of Contents

1 INTRODUCTION

The basic unit of the pixel detector is the module. A module is a rectangular active device approximately 6cm by 2cm with 47,268 pixels. The pixel detector is an array of nominally identical modules overlapped in precise relative positions to form continuous acceptance layers and disks. A detailed description of the integration of the modules into the full detector, a formidable task in itself, is beyond the scope of this document. Each module is composed of a bare module (the subject of this document), a flex hybrid, a module control chip (MCC), and a flex pigtail cable, as illustrated in Figure 1.

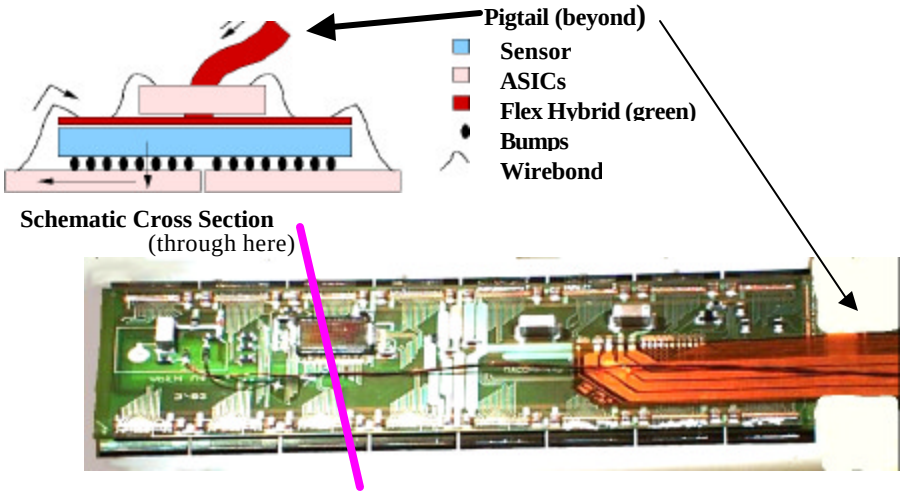

Figure 1: *schematic cross section of a pixel module (top) and photo of a prototype module with disk pigtail (bottom).*

The bare module is made of a silicon sensor tile (21.4 x 62.4 x 0.25 mm^3) and 16 front-end (FE) chips, each of 7.4 x 11.0 x ~0.2 mm^3. The active surface of the module is of 16.4 x 60.8 mm^2 and corresponds to 41984 pixels of 50 μm in azimuth x 400 μm parallel to the LHC beam, and 5284 pixels of 50 x 600 μm. The longer pixels are necessary to cover the gaps between adjacent FE chips. The module has 46080 channels (160 rows x 18 columns x 16 front-end chips). The number of channels is 2.5% smaller than the number of pixels because there is a 200 μm gap in between FE chips on opposite sides of the module, and to get full coverage we must connect the last 8 pixels in each row to only 4 channels. Thus on 5% of the surface the information has a two-fold ambiguity that we believe we can resolve off-line. The pixel geometry and connections in the regions where four FE chips meet are illustrated in Figure 2.

All bare modules are geometrically and electrically identical. However, two different technologies are used to interconnect the FE chips to the sensor tile (see below), and two separate manufacturers of sensor tiles are used. This was done in order to have technical redundancy during the development, but it was also necessary to use multiple sources in order to fill the needed quantities.

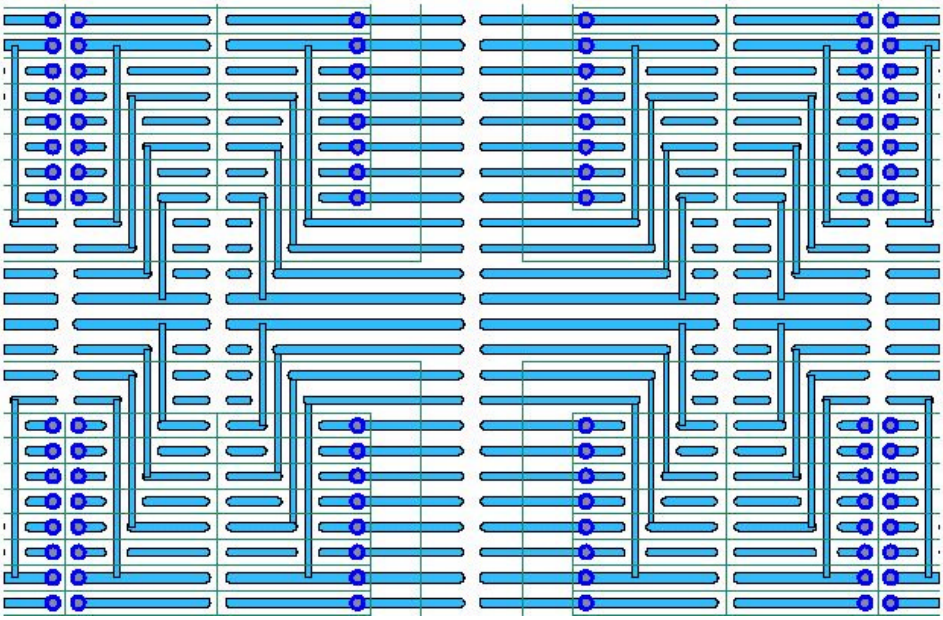

Figure 2: *the pixel geometry and connections in the region in between 4 chips. The thin lines indicate the boundary of each pixel circuit and the edges of the front-end chips.*

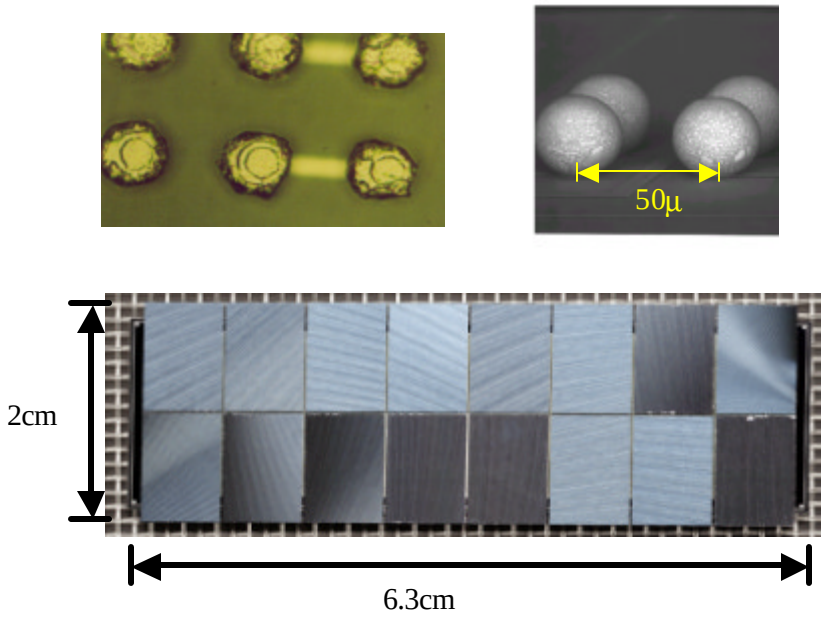

Figure 3: *Indium bumps (top left), Solder bumps (top right) and back side of bare module.*

The connection between each pixel and its read-out channel is made through a bump bond. Two technologies are used, indium bumps and solder bumps. Figure 3 shows a detail of both types of bumps as well as a bonded assembly (bare module). The minimum bump spacing is 50µm. In order to reduce material and module thickness (which allows for a tight geometrical arrangement), the FE chip wafers are thinned to 200µm.

No underfill material is used between the bumps to minimize the capacitive coupling between pixels as well as the capacitive load on the FE inputs. Consequently the bumped assembly is mechanically held together only by the bumps, which makes it vulnerable to stresses from thermal expansion mismatches. Glue interfaces between the flex and the bare module and between the module and the support structure must therefore be compliant. Low shear strength glues and special glue deposition patterns are used to achieve this. These interfaces are included in this document but will be fully described and characterized as part of a separate document on module assembly (ATL-IP-AN-0003).

2 REQUIREMENTS

2.1 PHYSICS PERFORMANCE

The physics performance requirements for the pixel bare module should assure:

1. excellent pattern recognition in high multiplicity environment
2. excellent transverse impact parameter resolution
3. good 3D-vertexing capability
4. excellent b-tagging capabilities

These performance requirements lead to the following design requirements:

1. minimum efficiency per pixel, more than 97 %, which should include dead pixels and masked noisy pixels
2. minimum efficiency per chip, at least 99 % for b-layer and 98 % for layers 1,2 and disks
3. minimum efficiency per module, at least 99.5 % for b-layer and 99 % for layers 1,2 and disks
4. the smallest practical size of the pixel, which is currently set by the electronic design 50 micron x 400 microns
5. minimum material in all elements, consistent with a realistic assembly yield and safe operation
6. the fraction of noisy pixels less than 10^{-5}
7. the charge cross-talk between pixels less than 10%
8. the threshold of 3000-3750 electrons with the dispersion less than 200 electrons after calibration
9. Equivalent noise charge <300 electrons
10. transverse impact parameter resolution with 11 micron constant term

2.2 ELECTRICAL PERFORMANCE

In order to determine the threshold and noise of each individual pixel, the amount of charge which is injected is scanned over a range of 0 to O(10,000e-). For each value of charge within the scan, 100 strobes are issued. As the charge magnitude passes the discriminator threshold the pixels start to generate hits, eventually reaching a plateau corresponding to the number of given strobes. A histogram of occupancy versus charge is built up for each pixel. The shape of this histogram arises from the fact that the noise present in the pixel channel

causes the threshold measurement to statistically vary in a Gaussian manner. This Gaussian distribution is integrated in this type of measurement giving rise to an 's-shaped' profile. Since there is no functional form for the integral of a Gaussian, an approximate error function is used in order to fit to the data and derive the threshold (given by the median) and the equivalent noise charge (ENC) which is given by sigma.

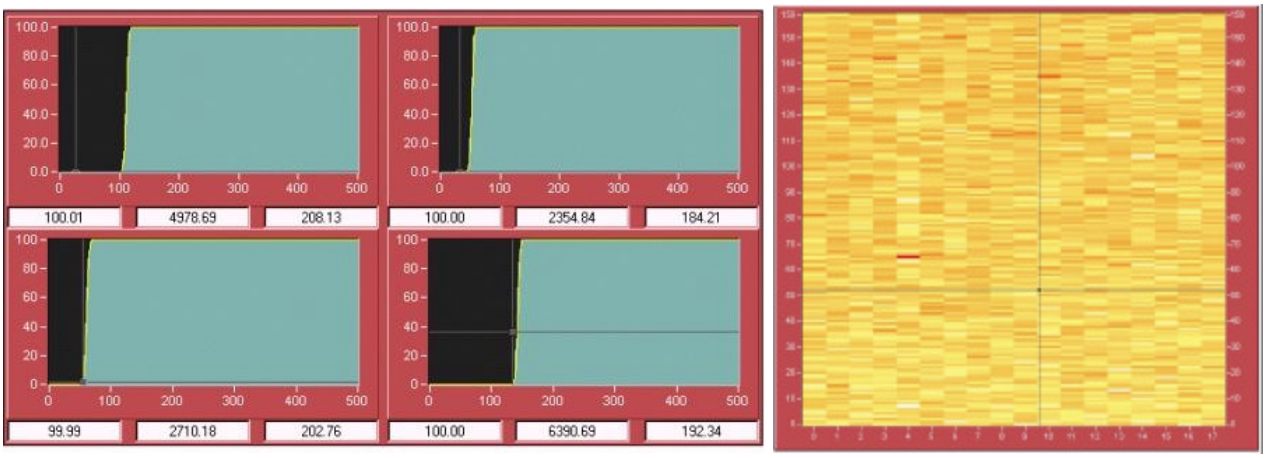

Figure 4: *S-curve histograms for a module threshold scan*

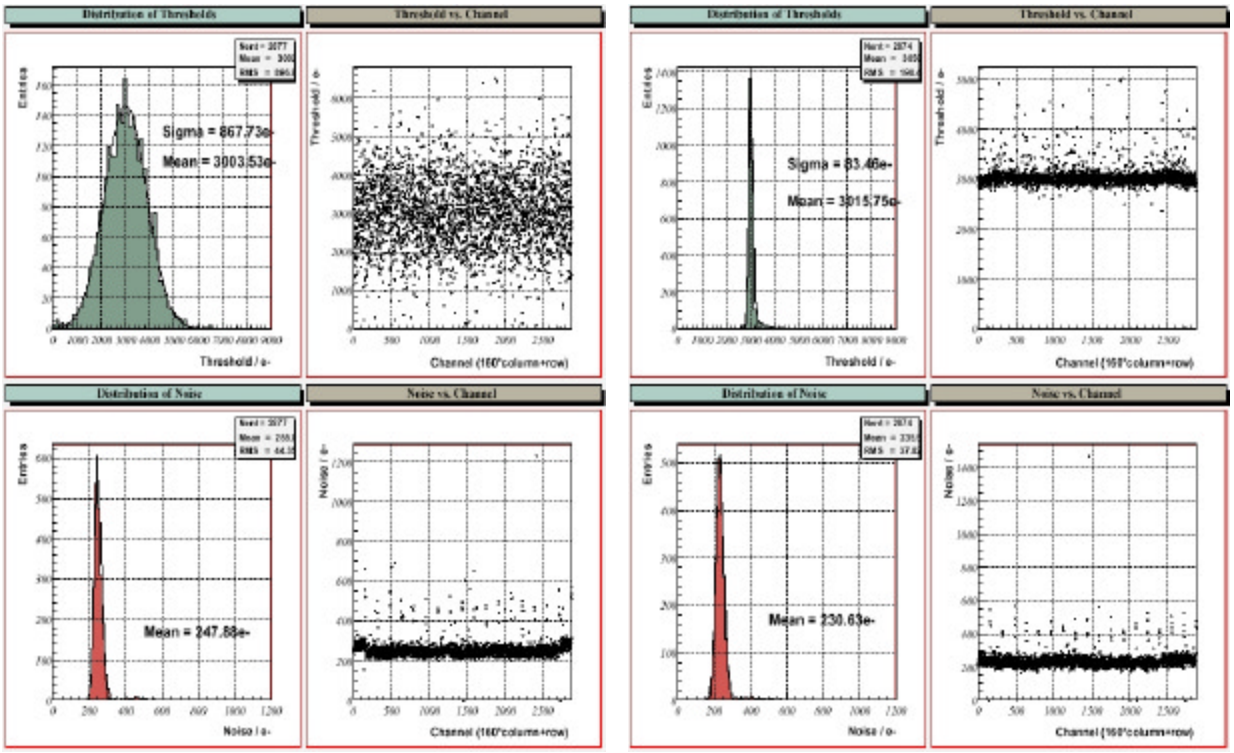

Figure 5: ***Threshold and ENC distributions for a module scan. Before tuning (left) and after tuning (right).***

Figure 4 illustrates four example s-curve histograms from a threshold scan performed on an FE-I1 chip. On the right of the figure the integral of all hits in these histograms is plotted as a geographical colour-scale map for a whole FE-chip, column number on the horizontal axis and row number in the vertical. The colour variation indicates the degree of threshold dispersion over the chip for a case in which the individual threshold trim DAC

settings have not been optimised in order to minimise the width of the threshold distribution. Note that every pixel in the chip is responsive, this is generally the case for FE chips which pass the most fundamental selection criteria at the wafer probe stage.

In order to meet the required performance demands in terms of fake occupancy and efficiency in ATLAS, the thresholds in all pixels need to be matched at the level of ~100e-. The FE-I1 pixel cell design incorporates a 5-bit threshold tuning DAC (TDAC) which provide a mechanism for making small relative threshold adjustments at the single channel level. Figure 5 shows some example threshold and ENC distributions for a single chip assembly in which a special single-FE-scale production-style sensor is bump-bonded to an FE-I1 chip. A Gaussian fit to the initial untuned dispersion has a sigma of 868e- a figure which is reduced to 83e- after tuning. The post-tune distribution has a slight upper tail but no channels at thresholds too low which would cause them to be inoperably noisy. The RMS of this distribution is ~100e-. In the lower half of this figure are the ENC distributions corresponding to the untuned and tuned cases. Before tuning the noise distribution peaks at 248e-, this comes down to 231e- after tuning since there are no longer any pixels at extremely low thresholds (in an oscillation condition e.g.) to influence the overall noise level of the chip.

Figure 5 shows threshold and ENC distributions for an entire 16-chip MCM. With careful TDAC tuning a threshold dispersion not dissimilar to the single-chip-assembly case (113e-) is achieved. The noise evaluation reveals an ENC of 263e- in the tuned state which also compares very favourably with assemblies constructed from single FEs.

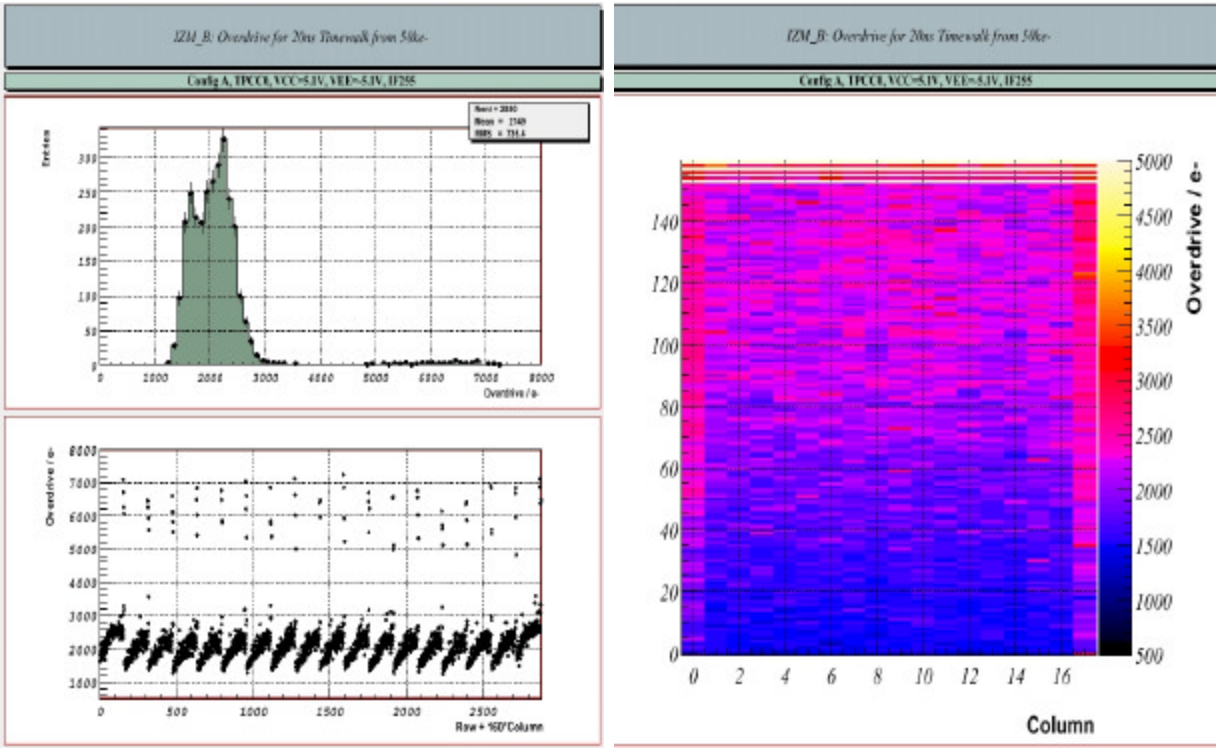

Figure 6: *In time overdrive distributions*

Critical to the efficient operation of the ATLAS Pixel Tracker is the ability to properly associate hits with their originating beam interactions. For smaller charges the analogue chain in the pixel cell takes longer to respond. This tends to be dominated by the preamplifier stage when a finite load capacitance, (i.e. from the

sensor), is present. In the absence of any capacitive load, the discriminator speed tends to be the limiting factor. The convention adopted here is to express this timewalk as the amount of overdrive (or charge-above-threshold) for which the discriminator fires 20ns more slowly than for an overdrive of 50,000e-. The choice of 20ns is made in order to allow some contingency for other sources of timing uncertainties, (e.g. jitter in the trigger distribution).

The strategy for evaluating timewalk in FE-I1 is to determine the relative response time for a large range of input charges. For each charge the level-1 trigger delay is set in order to be slightly too late (e.g. by 1 BCO) and the precise delay of the calibration hit-strobe is scanned with respect to the trigger. As this delay is increased (in 0.66ns steps) the hit is gradually pushed forward in time until it eventually matches the trigger and is read out of the chip. The resultant histograms of occupancy versus strobe-delay essentially have the appearance of a step function, except the step has a finite width due to the projection of noise in the channel onto the time axis, (as per the derivative of the timewalk function). Fitting an error function to such a histogram yields an accurate relative time measurement which is given by the error-function median. The timewalk function is derived by plotting these median points versus the difference of the input charge and the known threshold for each pixel. A process of interpolation is then used to find the 50,000e- overdrive point and the 20ns timewalk point in order to extract the minimum in-time overdrive limit.

Figure 6 shows the distribution of in-time overdrive for an FE-I1 single-chip assembly as a histogram, a scatter plot (overdrive versus channel-ID) and a colour-scale map representation. Clearly there is a systematic worsening of the timewalk with increasing row number which is due to a deficiency in the distribution of the main preamplifier bias current (IP). This issue is being addressed in FE-I2. The overdrive values for the very lowest row numbers are the most relevant since those channels are in receipt of the correct biases. For these channels an overdrive of ~1300e- is recorded indicating an `in-time threshold' of 4300e- for the nominal global threshold of 3000e-.

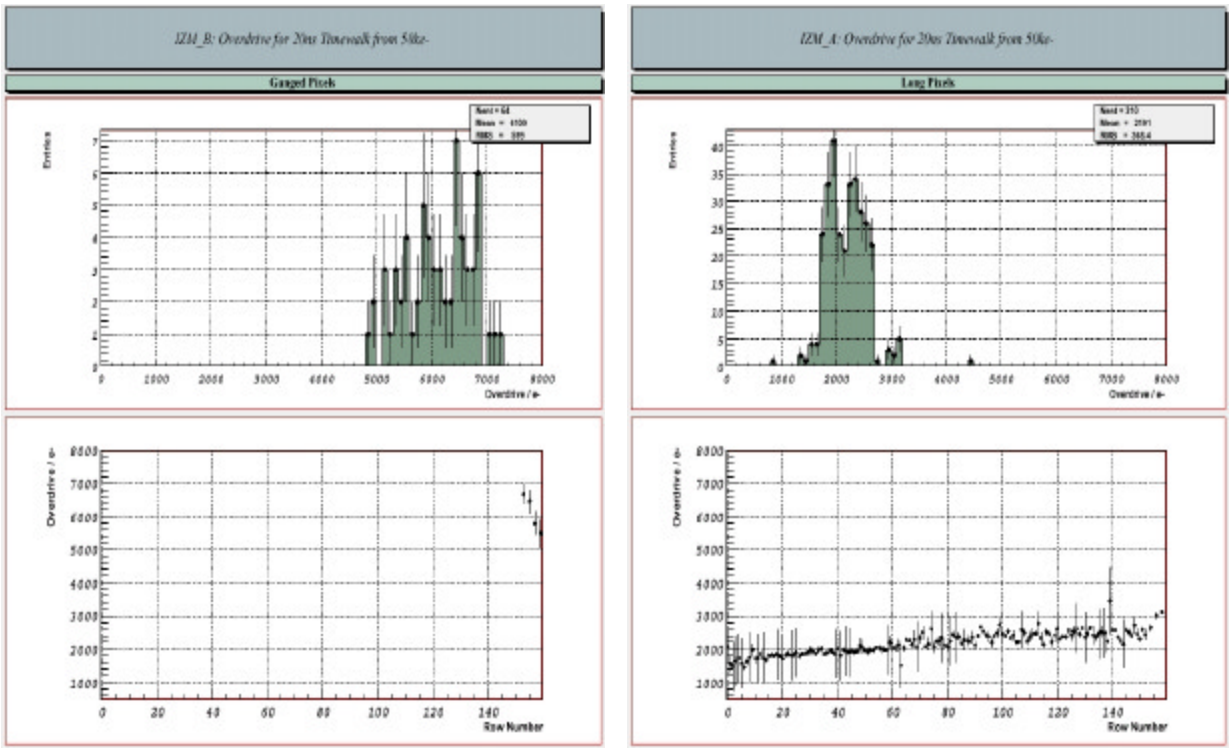

Figure 7: ***Long and ganged pixel distributions***

In Figure 7 the ganged and long pixels are individually examined. Since the ganged pixels are at the top of the chip the fact that the load capacitance is much higher is exacerbated by the poor preamplifier bias distribution. An in-time overdrive range of 5ke- to 7ke- is measured. For the long pixels the figures are between 1.7ke- to 2.7ke-. In FE-I2 provision is being made to supply the ganged pixels with a much enhanced preamplifier bias current to compensate for their high load. Since they only account for a small percentage of channels, the impact on the power budget in taking this measure is minimal.

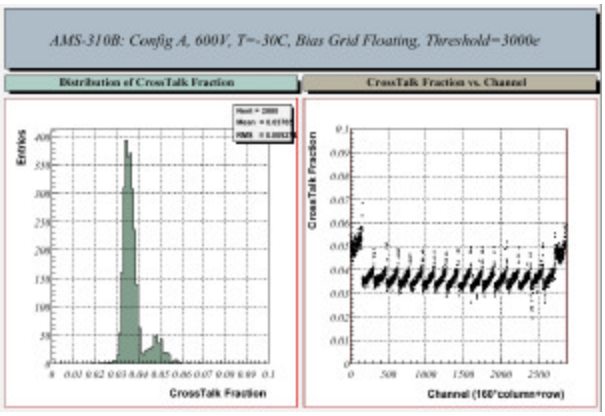

Figure 8: ***Crosstalk measurement***

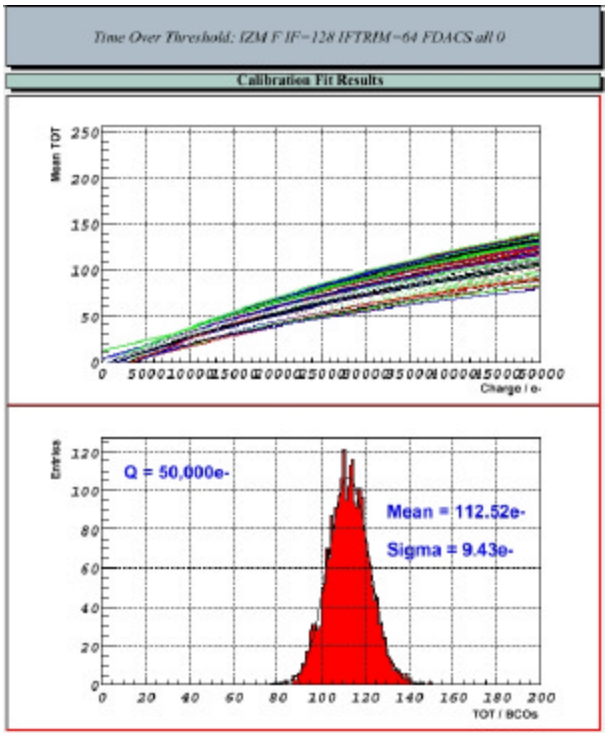

Figure 9: ***Relationship between TOT and charge***

The method for determining the degree of charge loss to neighbouring pixels (i.e. the analogue crosstalk) involves enabling a pixel for which the threshold is known and injecting a range of charge into its neighbours up to very high (~200ke-) values. As the magnitude of the injected charge is increased, eventually the degree of charge which couples into the readout-enabled pixel is sufficient to cause its discriminator to fire. The percentage of

crosstalk is then simply evaluated as the quotient of the threshold and the median charge for this to occur (by fitting an error-function). Figure 8 shows the distribution of crosstalk for an FE-I1 assembly. For the regular 50um X 400um pixels the crosstalk is determined to be 2.4% while for the special channels (ganged pixels and long pixels) the figure is 3.9%, comparing very favourably with the 10% requirement. In Figure 9 the relationship between the 8-bit Time-Over-Threshold (TOT) charge measurement and input calibration charge is expressed as a set of functional fits for an entire FE-I1 assembly and as a distribution of Mean TOT for an input charge of 50ke-. For these data no attempt has been made to match the TOT calibrations channel to channel by adjusting the 5-bit feedback current trim DACs. Without any tuning the matching is already better than 10% with a dispersion of 9.4BCOs recorded for a distribution which peaks at 112.5BCOs.

2.3 MECHANICAL PERFORMANCE & CONSTRUCTION TOLERANCE

The bare module is an intermediate assembly, which then will be dressed with the flex hybrid module giving the full module assembly. The maximum planarity error of the sensor is specified to 50 μm. The planarity of the full module assembly should be of the same order to allow a correct mounting on the local support surface. Therefore also the flatness requirement for the bare module assembly should be of the same order. Drawing number ATLAY0001_4 controls the envelope of the module assembly and its components, bare module assembly included. The FE electronics chips are mounted on the sensor with high alignment precision, so that the bare module assembly is well within the envelopes without requiring any additional provision in terms of mechanical alignment.

The minimum normal operating temperature of the pixel silicon sensors is –25°C, but during transient conditions the minimum temperature may reach –35°C. The design of the bare modules and interfaces to them will assume that no damage both structural (loss of integrity) and geometrical (loss of shape, distortions) occurs in case of single system failure resulting in partial or complete loss of coolant with power on (one hour T runaway once per lifetime). The bare modules as well as all other module parts and assemblies have to be qualified against temperature cycling, according to Table 1.

TRANSIENT MODE	TEMPERATURE CYCLE (°C)	NUMBER OF CYCLES (LIFETIME)
Single module On-Off	20	750 (*)
Cooling circuit On/Off	30	750
Detector warm-up	45	30

(*) It is expected to have one On/Off transient per day over the whole detector, the worst case for an individual module would be realistically to have 1 cycle per week.

Table 1: *Module temperature cycle requirements*

During the operation it is not expected to have any significant mechanical action on the bare module or any more severe mechanical action than those introduced by the tools during the module assembly and loading on the local supports. Therefore there is little risk of breaking the bare module assembly during operation. However the CTE mismatch between the bare module and the local support at one side and the flex hybrid circuit on the other side might induce, as function of the stiffness of the glue layers, some stress on the bumps. These stresses are negligible as single event, so they are not affecting both the sensor and the FE chips, but they might induce some fatigue on the bumps as results of the thermal cycling. The mechanical stresses are therefore driven by the thermal ones. The qualification of the bare module assembly against the mechanical actions is then enveloped by the qualification against thermal cycling. The bare module will be assembled in clean areas and then stored in dry N2. The ambient operating atmosphere around the bare module will be dry nitrogen with up to about one-part-

per thousand of C_3F_8. The construction of the bare module shall be such as to meet all requirements under such a conditio n.

2.4 GLUE INTERFACES (CLAUDIA)

3 BUMP BONDING ASSEMBLY

Production of a bump bonded assembly is a four-step process:

a) Bumps (or Under Bump Metalization) are deposited on sensor wafers and bumps are deposited on electronics wafers,
b) The electronics wafer is thinned down to 180 μm
c) The sensors and the electronics wafer are cut and sensor tiles and front-end chips are selected (we keep here only the known-good-die)
d) The front-end chips are flipped on the sensor tiles after precision alignment. The electrical and mechanical connection is formed at appropriate temperature and pressure.

The steps (a) and (d) depend on the technology, indium or solder. Both bump bonding and wafer thinning are techniques of increasing importance in consumer electronics products, which have continued to advance even as the pixel module development was in progress. Nevertheless, the fine pitch (50 μm) and high density of connection (~5000/cm^2) chosen for the ATLAS module design are not of immediate industrial interest and this greatly limits the number of possible vendors. In turn the small bumps required for such fine pitch are more vulnerable during wafer thinning than commonly used larger bumps. Thinning of the FE wafers per se is not a major problem, for even as the FE electronics wafer diameter has grown to 200mm the capabilities of vendors have evolved to match.

3.1 INDIUM BUMPS

Indium bumps are obtained by electron beam evaporation of In under vacuum (~10^{-6} Torr) and subsequent deposition of the metal through a 15 μm thick polyimide layer spun on the wafer and conveniently patterned. The patterning of the photosensitive polyimide is obtained by etching after exposure to UV light through a high precision contact mask. Holes of appropriate size, position and profile allows the deposition of In cylinders of ~20 μm diameter and ~8 μm height in about 15 minutes after the vacuum is established. The polyimide (and the excess In) are then removed with a wet lift-off process. In order to have good adhesion of the deposited In to the pixel electrode, a layer of appropriate Under Bump Metal is deposited first and cleaned immediately before the deposition with a plasma etching process. The deposition process is illustrated in Figure 10. One 8” wafer is processed per cycle.

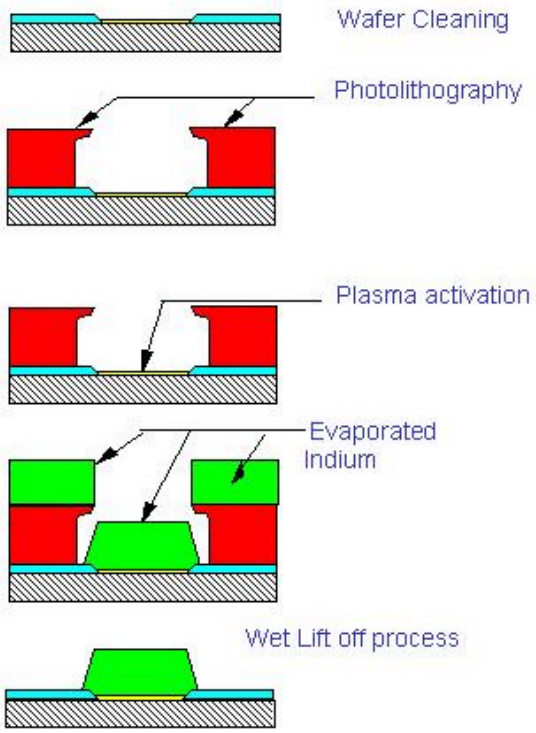

Figure 10: *Diagram of indium bump deposition sequence.*

3.2 SOLDER BUMPS

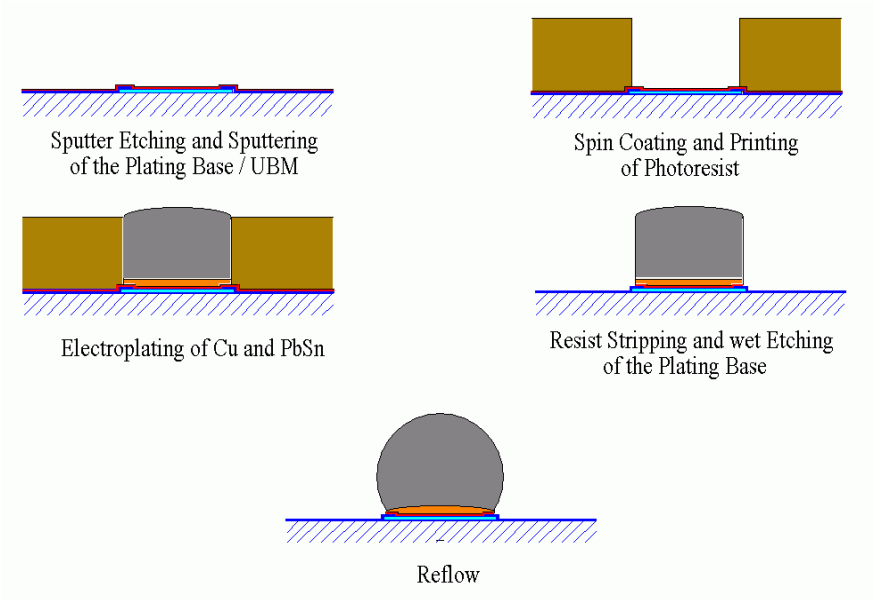

Figure 11: ***The galvanic solder bumping process on 8" IC wafers and the different steps of the bump deposition***.

Solder (Pb 40% Sn 60%) bumps on IC wafers and UBM (Under Bump Metallization) on sensor wafers are both fabricated by electroplating. The IC wafers are sputter etched before sputtering of an adhesion

layer (Ti,W) and a plating base (200 nm Cu). Photoresist is spin coated on and patterned with the bump locations. A copper layer (5 µm) is then plated as a wettable under bump metallization on which the Pb 40% Sn 60% solder is deposited. After photoresist removal, cylindrical bumps about 30 µm high remain which are turned into spherical bumps in a reflow process step. The bumps are ~25 µm in diameter. See Figure 12. The processing of the UBM on the sensor wafers is identical but ends with the 5µm copper layer. The processing requires spinning, etching, plating, and automated inspection equipment and is optimised for 8" wafers.

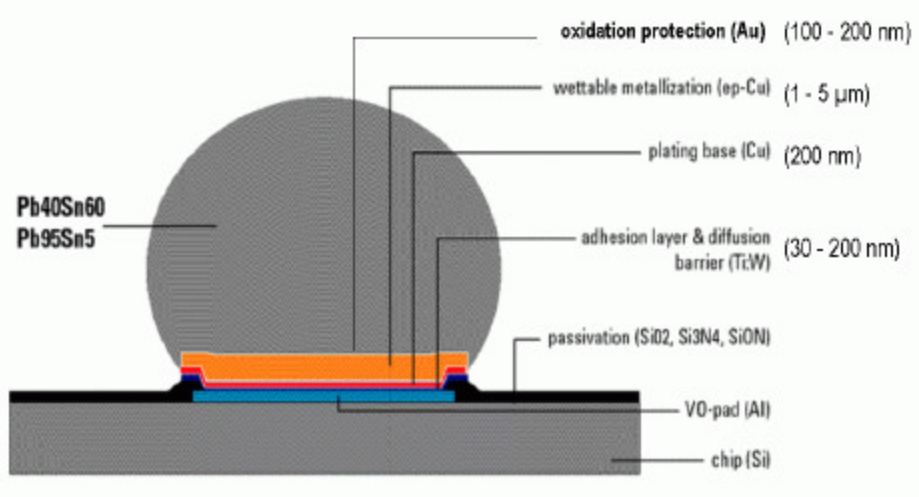

Figure 12: *Cross section of a solder bump*

3.3 WAFER THINNING AND SINGLE CHIP PROBING

Before the bumped FE chips are flipped onto the sensor tiles they must be thinned. Thinning takes place after bump deposition because the many processing steps to deposit bumps cannot be carried out on fragile, thinned wafers. After thinning the wafers are diced and the individual die are electrically probed. This is to ensure that KGD (known good die) are used to build bare modules, for even though the wafers had been probed before bumping, the processing, thinning, and dicing can potentially damage some of the chips.

The wafers (both solder and indium) are coated with thick photoresist completely covering and protecting the bumps. Thin (~180 µm) wafers are obtained by backside mechanical grinding. This is done by companies specialized in wafer thinning. The thinned wafers are immediately diced, as they are very fragile and could otherwise easily break.

Two methods have been devised to test diced and bumped ICs immediately before flip-chip. Both methods are based on custom carriers that can hold many die on a probe station. Both methods have been shown to work and both will be used (at different sites).

3.3.1 CHIP PROBING METHOD 1

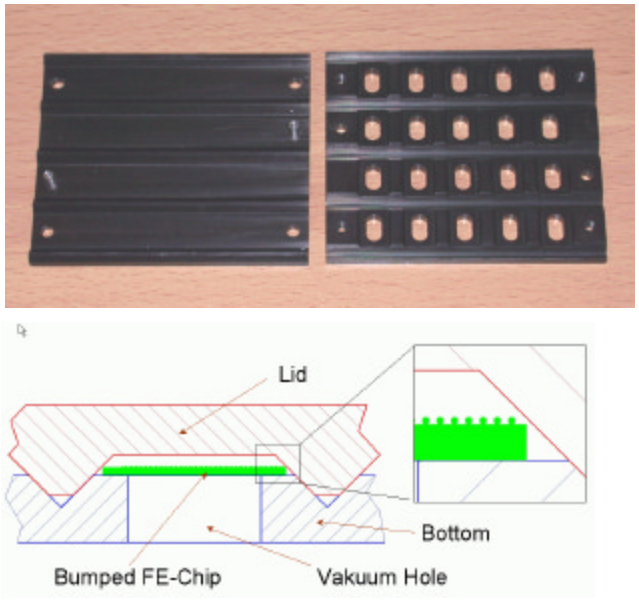

Figure 13: ***Method 1 chip carrier. Photo (top) and cross section (bottom).***

A special carrier for bumped ICs (see figure) has been designed, tested and fabricated in some numbers (see **Error! Reference source not found.**). More carriers are in production. 20 ICs can be placed in a carrier. The carrier has a lid that holds die in place by contact with the edges only and therefore avoids damage to the bumps. This lid has opening that allow access to the IC bonding pads for probing, but access through the lid requires a special "deep access" probe card. The ICs are tested without the need to take them out of the carrier. The same full suite of electrical tests carried out of full wafers is repeated here. The testing process uses an automatic probe station with pattern recognition capability which renders the precise alignment of ICs under the probe card unnecessary. With this carrier it is also possible to load the die in the carrier still coated with photoresist and then wash the resist off for the entire carrier.

3.3.2 CHIP PROBING METHOD 2

This method was developed and successfully used for probing 10,000 SVX3 readout chips for the CDF Run 2 silicon strip detectors, presently taking data. Those chips did not have bumps but the method works equally well for bumped chips. The carrier for this method has precision cavities (formed by a laser cut steel plate) to register the die in a well defined angular position (see Figure 14). The chips are registered by simply pushing against two edges. Once registered the carrier has vacuum hold-down for each die. The carrier loaded with chips is placed on a probe station and the control software cycles once through all the die locations without lowering the probe needles. An operator must either accept or correct each die location (correction is done by moving a joystick). This process is very fast- only a few second per die. Once the program has stepped over all die in this way the correct location of each die has been memorized and the carrier can be automatically probed like a regular wafer. Neither a probe station with pattern recognition nor a special probe card are required. The chip carrier has a transparent lid used to protect the bumps while probing is not in progress. With the lid on, it is possible to photograph the bumps on the probe station. A program has been developed that takes pictures of all bumps on every die and automatically record the bump row and column for each picture. The purpose of this is

both to monitor bump damage during thinning, dicing, and handling, and to have a visual record of bump defects that can be correlated to electrical anomalies seen later when operating a module.

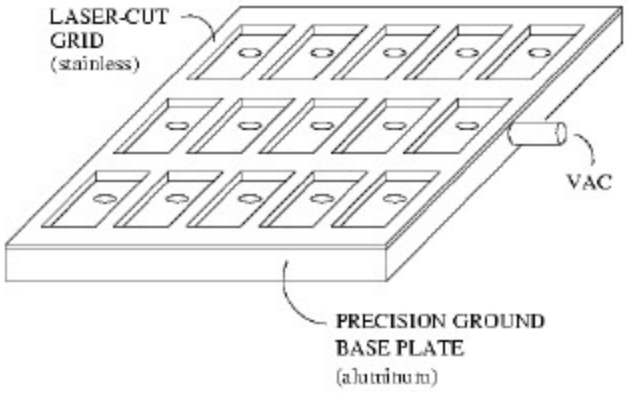

Figure 14: *Conceptual drawing of method 2 chip tray.*

3.4 FLIP-CHIP OF BUMPED PARTS

Sixteen bumped and thinned ICs (after dicing and electrical probing) are flipped to a sensor tile to make a bare module. For Indium the ICs with bumps are flipped and bonded one at a time onto the sensor, which also has bumps. The flip chip process involves precise alignment of each chip to the sensor. The bonding is a thermo compression process at 90 C, that takes a couple of minutes per chip (once aligned). All the operations (alignment, placement, bonding) are done using one machine.

The solder bumped ICs are flipped to sensor tiles with UBM and subjected to a soldering process. The ICs are tacked to the sensor UBM by solder flux using a precision pick and place bonder. The entire module is then placed in a reflow oven in which the bumps and the UBM are solder merged in a heating cycle of 4 minutes with a maximum temperature of 240^0 C for a few seconds in an activated atmosphere. Reworking of badly flipped ICs by detaching a chip and replacing another one has been successfully demonstrated.

4 DESIGN & PROCESS VALIDATION

4.1 PHYSICS SIMULATION

Recent simulations have been done on the sample of signal pp->WH->lnubb events with mH=100 GeV and mH=400 GeV with low luminosity pile-up L=2 10^33 cm^-2s^-1. For the background events the reaction pp->WH->lnuuu was used with the same values of Higgs mass. In case of the full 3-layer pixel system with 400 microns long pixels and using 3D b-tagging method the u-jet rejection with 60% b-jet efficiency is Ru=106 in case of mH=400 GeV and Ru=119 in case of mH=100 GeV.In case if 300 micron pixel modules are used in the b-layer the rejection is improved by ~10% for mH=400 GeV. These results are for a perfect (defect-free) detector. The loss of 1% efficiency of chips and modules results in the deterioration of the u-jet rejection by 5-7%. So all efforts should be taken to increase the efficiency of the modules. The modules with the best efficiency should be selected for b-layer.

4.2 BEAM TEST RESULTS

In Summer 2002, modules built with the first generation of rad-hard electronics and pre-production sensors have been tested in the H8 beam line of the SPS at CERN, using 180 GeV/c hadrons. The test beam setup consisted of a beam telescope, trigger scintillators and a cold box where pixel devices were placed. The telescope consists of four planes of double sided microstrip detectors, with a resolution of approximately 5 μm in both projections. Each plane has an active area of 38.4×38.4 mm^2. The trigger is given by the coincidence of two plastic scintillators in front of the telescope and a third placed after the last plane. For each trigger, a TDC measures the time difference between the scintillator pulses and the edge of the 40 MHz clock which is used to drive the pixel electronics. This timing information has a resolution of 36 ps and is used for time-walk and in-time efficiency measurements as explained below.

When a trigger is issued, the modules are requested to send out data collected in sixteen consecutive clock cycles around the trigger time. That results in an observation window of 400 ns instead of the operational one at the LHC of 25 ns. The wider readout time was used to get information about timewalk (section 4.2.2) and noise (sections 4.2.1 and 4.2.3).

The cold box was placed between the second and third microstrip planes. Cooling was not used for module data taking. The possibility to change the box position by remotely controlled step motors allowed an easy scan of modules. A VME based read-out system was able to record between 7000 and 10000 events for each SPS burst of 4.8 s. Data from a scan of three modules have been analysed in order to collect information about efficiency, uniformity of response and noise rate.

4.2.1 DEFECTIVE CHANNELS

A full module scan in the test beam allows the full mapping of defective channels, either because of high noise rate or bad bumping. A typical channel noise is well below 300 e^-, with a typical threshold during test beam operation of 3000 e^-, the threshold to noise ratio 10:1 assures the random noise to be at a negligible rate (see section 4.2.3 for further details). Fixed pattern noise, due to a small number of channels with high activity may be dominant. These pixels can be either set at a high threshold or disconnected from the readout, in both cases trading off efficiency to get low noise occupancy. These pixels can be identified looking both at an anomalous occupancy and at the fraction of hits which are not in time with the beam crossing. As far as bump defects are concerned:

- disconnected bumps are visible as channels with no hits, but which behave normally from the electronic point view;
- merged bumps show up as channels with no hits but in which one of the neighbours either is very noisy because of the increased channel capacitance or has a count rate higher than normal.

The identification of the bump defects therefore requires a cross check with the laboratory tests.

Table 2 shows a summary of the defects observed in the three tested module. The normalization takes into account only the number of working column pairs. For bump bonding defect characterization, ganged pixels were excluded since they show a general efficiency problem and therefore having no hit is not enough to imply the existence of a bump defect. All modules show a rate of defective bumps below 10^{-3}. Also the total rate of noisy channels is acceptable: their masking results in an efficiency loss $<10^{-3}$ for the modules with IZM bumps and of $<10^{-2}$ for the module with AMS. The AMS module had a higher level of noisy pixels because it was built

with chips qualified as good on wafer (i.e. prior to bumping, thinning and cutting), whereas the IZM modules were built with chips individually re-tested just before flipping. These results underscore the importance of chip testing prior to flip-chip to build modules with only known good die.

Device S/N	Sensor	Bumps	Electronics defects		Bump bonding defects	
			Dead FE	Dead and noisy channels	Disconnected	Merged
20210021410057	Tesla	AMS	2	294/37760	21/36816	0/36816
20210020410085	CIS	IZM	0	7/42880	6/41808	0/41808
20210020420001	CIS	IZM	0	36/43200	2/42120	2/42120

Table 2: *summary of defective channels for the test beam modules (see text for explanation of normalization factors).*

4.2.2 IN-TIME EFFICIENCY

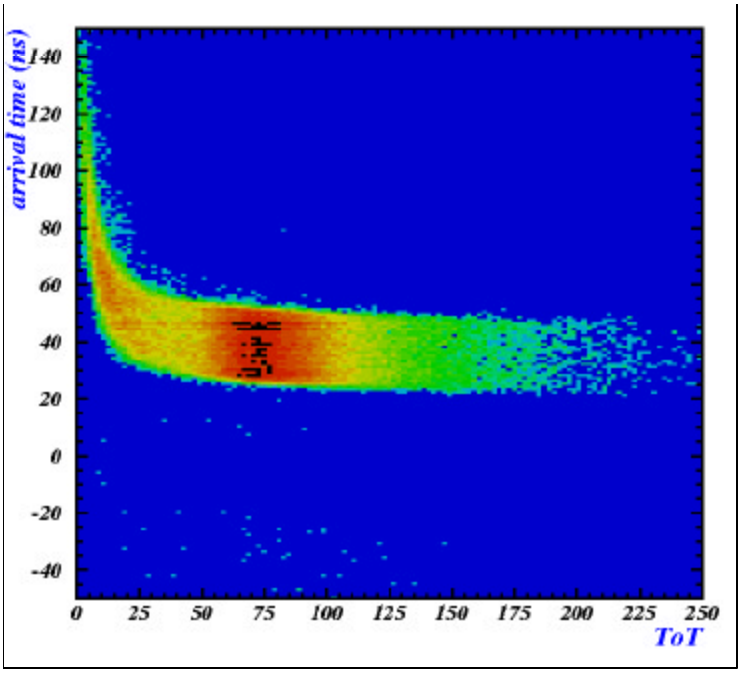

Figure 15: *Distribution of hit arrival times as a function of the ToT (in clock cycles)*

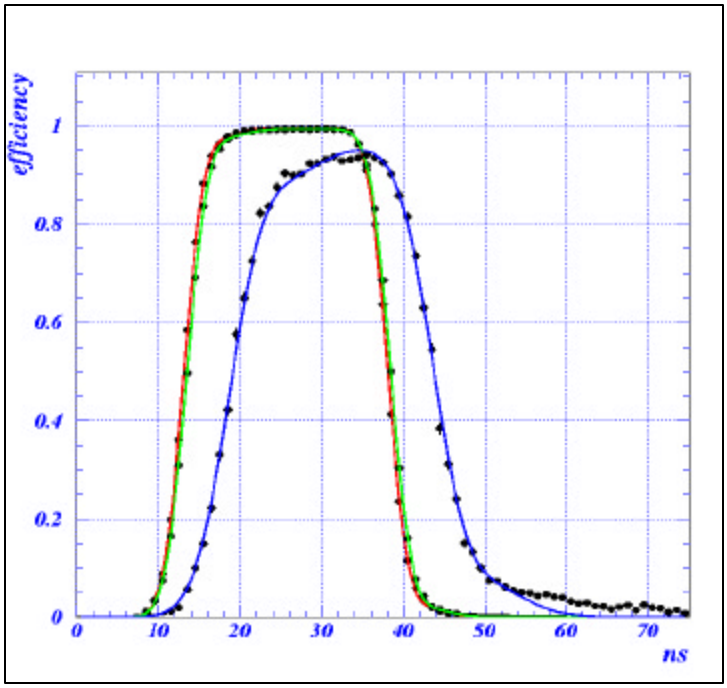

Figure 16: *efficiency vs. delay time curve for a run of module s/n 20210020410085, red line is for standard pixels, green for long pixels and blue for ganged ones.*

For the pixel detector operation at the LHC, particle hits must be associated to the correct beam crossing. Therefore the efficiency is not simply the probability to collect a hit when an ionizing particle crosses the detector, but must take into account the probability of misassigning the beam crossing. The timewalk, i.e. the variation in the time when the discriminator input goes above threshold, is an issue since hits with a low deposited charge have an arrival time later than the ones with high charges. The timewalk translates into a distribution of hit arrival times and the in-time efficiency is the integration of this distribution in a 25 ns window.

By construction the front-end electronics perform the integration over the 25 ns time window, and the TDC information defines, event by event, the time delay between the start of the integration and the crossing time. In-time efficiency can therefore be measured as a function of the delay. At the LHC the delay time will be tunable at the module level and can be set for maximum efficiency. Figure 15 shows the observed distribution of hit arrival times as a function of the ToT (Time over Threshold) which is used to estimate the deposited charge. In Figure 16 an example of efficiency vs. delay time curve is shown. In case of no timewalk effect it should be a square function with 25 ns width. The effect of timewalk is to smear this function. The observed value of the in-time efficiency have been fit with a parameterization which assumes the hit arrival time distribution is the convolution of a gaussian and an exponential distributions. This simple parameterization satisfactorily fits the data.

For standard 400×50 μm^2 pixels the efficiency for normal incidence particles is 99.57±0.15%, where the error is the r.m.s. of the observed value on the 46 working front-end chips. The 600×50 μm^2 pixels have overall performances similar to the standard ones. Ganged pixel have a stronger timewalk, resulting in efficiencies between 90% and 96% on 5% of the module active area. This issue will be addressed in the next version of the front-end electronics.

Serial number	Sensor	Bumps	Efficiency [%]	T spread [ns]	s [ns]	t [ns]	<ToT>/m.i.p. [clock cycles]
20210021410057	Tesla	AMS	99.59±0.14	1.6	3.16±0.14	1.07±0.05	76±5
20210020410085	CIS	IZM	99.60±0.20	1.4	2.65±0.10	0.88±0.09	87±4
20210020420001	CIS	IZM	99.52±0.10	1.7	2.66±0.18	0.89±0.08	79±5

Table 3: *summary of in-time efficiency data for the test beam modules (see text for explanation of errors).*

The position of the rising edge of the in-time efficiency curve can be used to measure the spread in timing among the different chips within one module. This is an important parameter since any chip-to-chip non-uniformity cannot be corrected after the module is built. Defining the time spread as the difference in timing between the earliest and the latest chip in one module, the highest observed value is 1.7 ns, which allows for a quite safe margin of operation. A summary of the in-time efficiency measurements for the analysed modules is displayed in table 2. Errors are the r.m.s. of the distributions within the module.

4.2.3 NOISE RATE

As mentioned in section 4.2.1, with a threshold/noise ratio of 10/1, random noise rate is expected to be fairly low. In principle noise can be measured selecting hits that cannot be correlated with the triggering particle because they are out of time and do not match the telescope extrapolation. Anyhow, most of the hits so selected are compatible in pulse height with a real particle signal. Also, given the H8 beam intensity, the observed rate is in qualitative agreement with what expected by a second beam particle passing in the detector within the 400 ns time window about the triggering particle.

When excluding this component, which gives an occupancy of 10^{-7} hit/pixel/event, the random occupancy rate per pixel per event is of the order of 10^{-8} for pixels in the region covered by the beam and one order of magnitude less in the region away from the beam. Given the spatial correlation of the noise and beam position, at present it is not possible to disentangle between residual background or noise somehow correlated to the activity in the front-end electronics.

4.3 THERMAL CYCLING AND DUMMY MODULE PROGRAM (TOBIAS)

4.4 PRE-PRODUCTION MODULES AND YIELDS

Pixel module prototyping and pre-production has been done with two firms: AMS and IZM, the first one using Indium the second one using solder (PbSn). Both techniques meet our specifications and we could not see any reason to prefer one or the other. In addition, using both vendors, we may more easily meet the production rate required by the ATLAS schedule.

The total number of modules built to-date (including so called daisy chain dummy modules and modules using predecessors of the ATLAS pixel FE-chip) is about ~60 (~ 35 IZM, ~25 AMS). The “hot” modules were built mostly with two types of electronics: FEB, a rad-soft pixel matrix developed in 0.8 µm CMOS

technology and FE-I1, a 0.25 μm technology version, rad-hard version of the same matrix with "close to production" characteristics. The two versions of the electronics chips are "bump-to-bump" compatible, i.e. they can be considered identical regarding the bumping and flip chip process. About half of the modules have been built using electronics chips thinned in the range of 150 to 225 μm. An analysis performed in the year 2000 on 3 modules (2 with thinned electronics) totalling ~10^5 bumps is documented in:

http://www.ge.infn.it/ATLAS/PixelWeek/Presentations/00-06_CG_BumpYield/00-06_CG_BumpYield.pdf

and indicates that, in those assemblies, the probability for short circuits is ~10^{-3} while the probability for a missing connection is ~2 10^{-5}, both numbers are well within the ATLAS specifications. Note that a short between 2 bumps results in a hit in one of the shorted pixels when a particle passes through either of the cells. The result therefore is a worse spatial resolution but no inefficiency.

Another 30 modules have been built with connection chains closed by the bumps. These also confirm the fault rate given above and did also allow to measure the bump-bond electrical resistance and the mechanical characteristics (e.g. force necessary to detach a bump bond). The bump resistance is ~ mΩ for the solder bumps and of ~10Ω for the indium bumps. The native oxide layer always present in the In bumps would initially set the resistance to ~100kOhm, but this layer is immediately broken when applying few hundred millivolts across the bump, which naturally happens through the input stage of the front-end read-out.

The force necessary to detach a bump is smaller for Indium, this case has then be better studied and a detachment force in excess of 0.1 g/bump has been measured. Calculations and mechanical tests on Indium based modules (including full thermal cycles) did not show this to be a problem. A larger series (50+50) of FEI1 modules is now in fabrication and this has been done in "close to production mode", namely:

a. Use KGDs both for sensors and for electronics
b. Use thinned electronic chips (and we test the dies prior to flipping)
c. Keep track of the module production steps through the Production Data Base
d. Provide micro radiographic data for all modules

The analysis of this pre-production (that will partially be presented at the PRR) should allow to finally set the limits for the production acceptance criteria and the ways to control them. The limits on number of dead or misbehaving channels due to the bumping and flipping operations should not sensibly impact the detector efficiency (i.e should be at a few per mille level) , which will be the case if the numbers found in the year 2000 will be confirmed.

5 QUALITY ASSURANCE

5.1 X-RAY INSPECTION

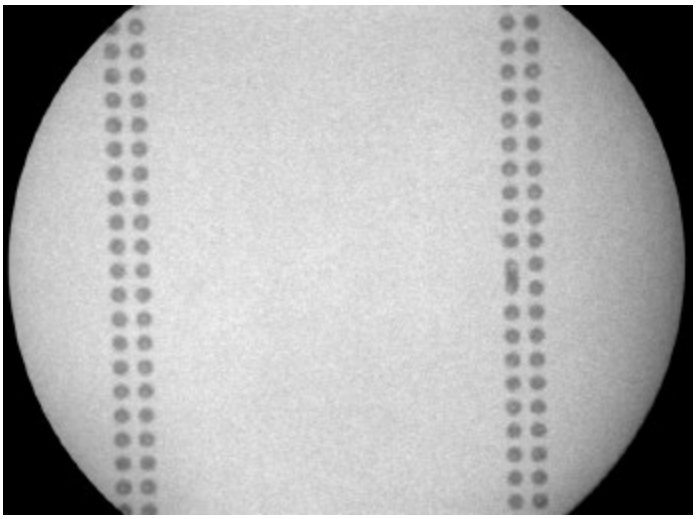

Figure 17: *X-ray image of section of solder bumped assembly*

High-resolution X-ray inspection has been used to evaluate the quality of the flip-chip process after production of bare modules. X-ray inspection of prototypes fabricated by all bump bonding vendors has been done and this inspection step will be used during the production bump bonding process.
X-ray inspection is a wide spread industrial tool for evaluation of printed circuit boards, flip chip assembly and other items. Very briefly, the sample is located on a low mass holder, illuminated by X-rays and viewed in real-time with a phosphor/CCD camera system. Depending on the vendor of the X-ray equipment, image averaging, enhancement, capture, etc is available via a software-controlled interface. In principle, some pattern recognition capability is also possible, but to date has not been used by ATLAS. All inspection is done by the operator in real-time. The resolution of the X-ray systems used is 2-5 microns, which is entirely adequate for quality control of bare modules.

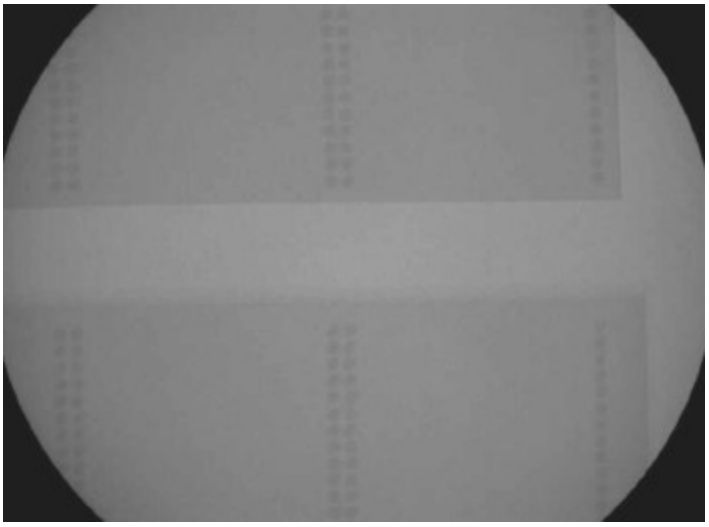

Figure 18: ***X-ray image of section if indium bumped assembly***

Figure 17 shows an assembly with solder bumps (a short between two channels is visible). The contrast of the image obtained with indium bumps is worse (Figure 18), but still good enough to inspect the connection quality after assembly. The x-ray inspection, which takes <1hr per module, allows to easily detect short circuits between bumps and, in the case of In, if pressure has been applied non uniformly over the chip in the flipping process. This may result into larger cross-section (but smaller height) bumps which may merge with neighbours in the location where pressure was higher. The detection of missing contacts can, on the contrary, only be inferred and a definite conclusion may only come from electrical analysis (noise of an unconnected front end is about half the noise of a connected front-end) or source scans. Anomalous small bumps indicate the possibility of missing connections.

5.2 ELECTRICAL PROBING

Basic electrical tests at the module level can be performed immediately after the flip-chip process. A more detailed characterization of the module behaviour can be performed after the full assembly. Bare module electrical tests can be performed by probing one FE chip at a time using the same probing tools described in section 3.3. Only the digital functionality test (section 5.2.1) will be performed, in order to detect front-end IC damaged after the flip-chip. Modules with one or more defective ICs will be rejected. Since testing of the assembled modules is much easier than probing, these tests will be useful in the pre-production phase, during the fine tuning of the process and yield assessment, and to allow re-working of a module (i.e. the replacement of a defective chip), which will not be possible after the mounting of the flex hybrid. Feasibility of reworking bare modules has been demonstrated for both In and SnPb bumps. At present, reworking is not expected to be necessary, but development is still ongoing to provide production tools for reworking in case the flip-chip yield is not enough to provide a sufficient number of good modules.

Full characterization of assembled modules will consist of the check of digital functionality, check of analog functionality, tuning of thresholds, ToT calibrations and mapping of defective channels. The goal is to use in the final detector only modules where the total fraction of defective channels is below 1%. All test results and the map of defective channels will be stored in the production database. In the following subsections the specifications of the above mentioned tests are given.

5.2.1 DIGITAL FUNCTIONALITY TEST

This test can be performed without applying the depletion voltage to a module. It consists in the writing and read back of all the configuration registers of the front-end IC, and of injection of pulsed at the discriminator output to check the readout of all pixel cells. Usually an IC either fails the configuration register test or is fully working. In the rare case of single cells not responding to the injection, they are added to the list of defective channels.

5.2.2 ANALOG FUNCTIONALITY TEST

An analog pulse corresponding to a charge of 20ke- is injected at the preamplifier input of each pixel cell. It should be above threshold for the default settings of the IC configuration. Pixels which do not react to charge injection are classified as defective channels. These can be either defects in the electronics or shorted bumps.

5.2.3 THRESHOLD TUNING

The thresholds can be tuned using a programmable DAC in each pixel cell. Threshold measurement is performed monitoring the answer of each pixel to a scan of injected charge. The injected charge for which 50% efficiency is obtained is defined as the threshold and the slope of the threshold curve provides the noise value. Thresholds must be tuned to 3000 e-, with an r.m.s.<200e- on the whole module. Average noise value on the module should be <300e-.

Channels with anomalous noise adjacent to channels failing the analog functionality tests are candidate for shorted bumps. Channels with anomalous low noise are candidate for disconnected bumps. That can be verified by redoing a threshold scan without applying the depletion voltage to a module. That is equivalent to short all bumped pixels together, while not-bumped ones still behaves properly from an electrical point of view.

5.2.4 TOT CALIBRATIONS

The relationship between TOT and charge is obtained monitoring the TOT response to pulses between 5000 and 50000 e-. Since a hit is readout only after the discriminator input has gone below threshold, one must ensure the TOT is small enough to have most of the hits going below threshold before the trigger latency has been elapsed. Test beam efficiencies have been obtained with an average TOT value for a minimum ionizing particle is equal to 40% of the maximum latency. That means the feedback current must be tuned to have a return to baseline within 40 clock cycles for a pulse of 25000 e-.

5.2.5 MAPPING OF DEFECTIVE CHANNELS

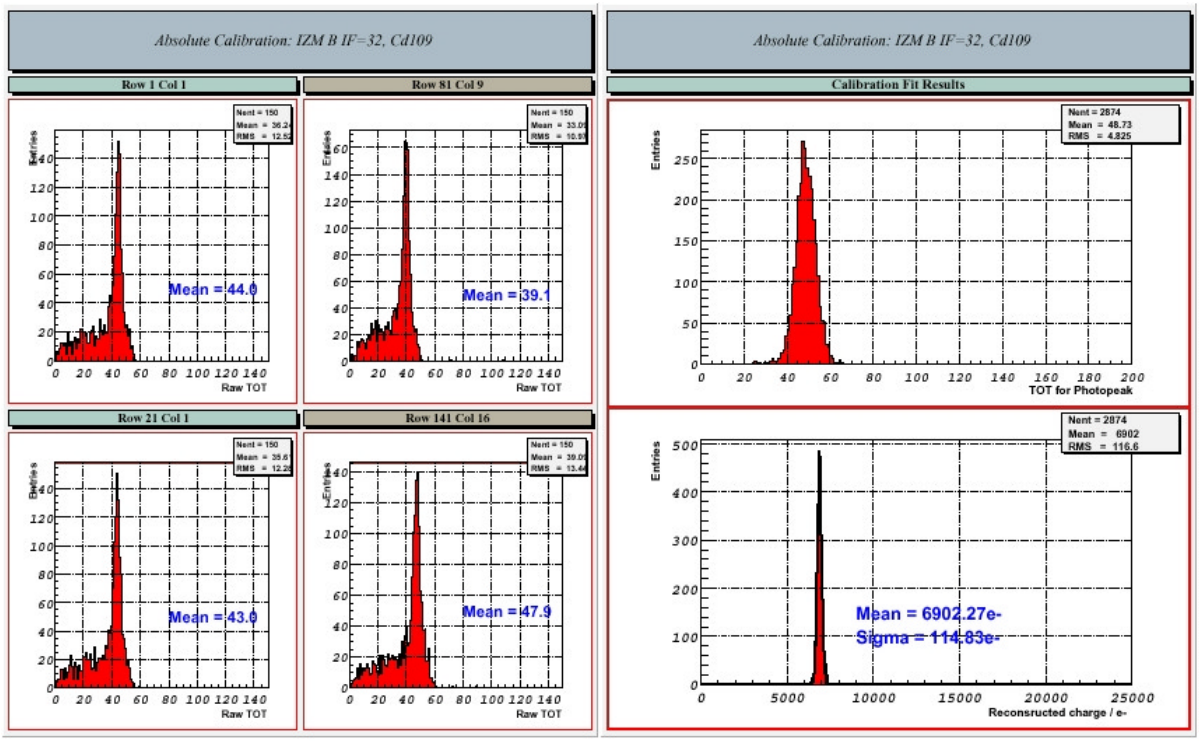

Figure 19: *Cd137 source data as explained in the text*

A list of defective channels is already provided by tests 5.2.1, 5.2.2 and 5.2.3. As an additional step, fixed pattern noise can be checked with random triggers and all channels with an occupancy greater then 10^{-4} must be masked and classified as defective.

5.3 RADIOACTIVE SOURCE SCANS

FE-I1 incorporates self-triggering circuitry which negates the requirement of providing triggers to the Pixel module externally. This is particularly valuable when testing modules using X-ray sources such as Cd_{109} and Am_{241}. Within each FE chip, a common OR of all selected discriminators serves to initiate the generation of an internal level-1 trigger following a fixed 64 BCO delay. For each event the FEs must first be 'armed' by providing them with a pulse on the external level-1 input. This is generated by sending trigger commands to the MCC. Hits which match the internally-generated level-1 trigger timestamp (according to a programmable latency period) are automatically read out of the FE. The MCC event-building architecture relies on serial FE data being registered in all of the enabled receiver channels, before the data for the whole event may be transmitted. For this reason the most efficient technique for acquiring source data for a whole module is to enable four neighbouring FEs at a time and to direct the source at common corners of those FEs.

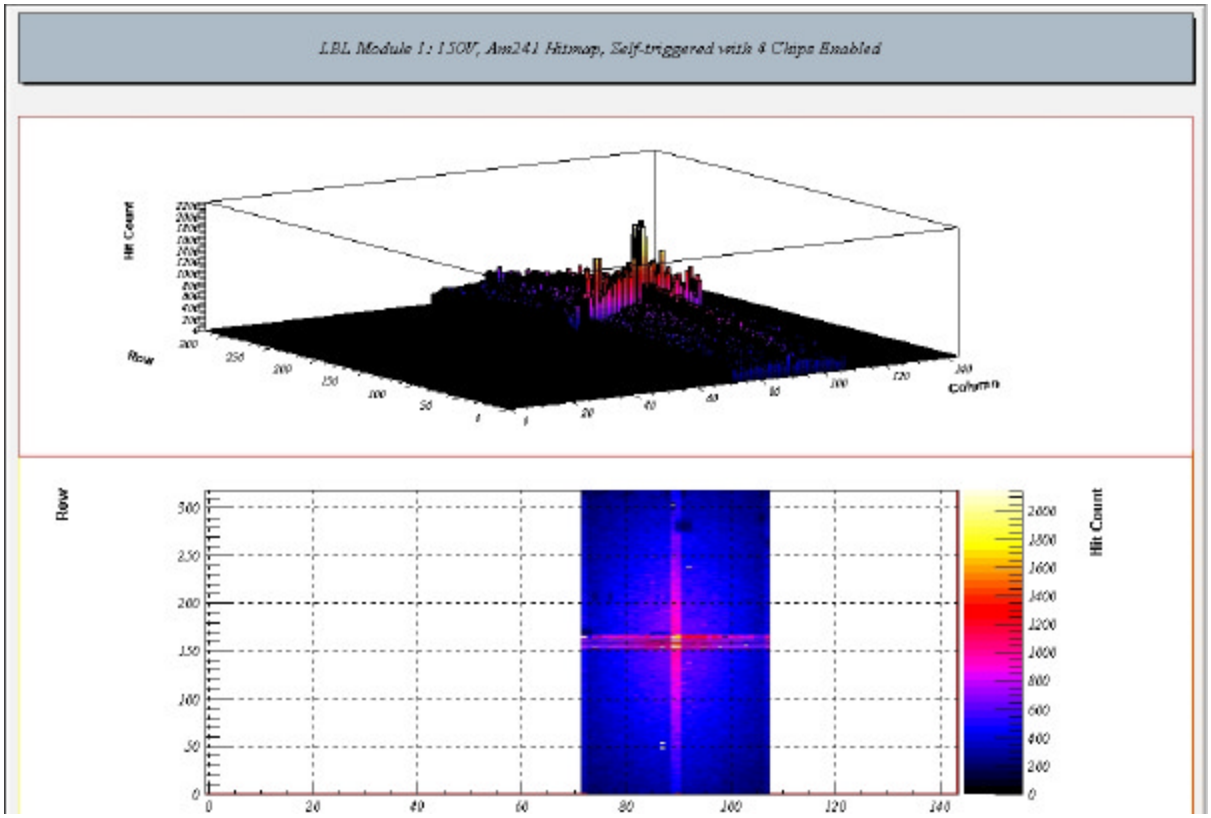

Figure 20: ***Am241 source scan***

The left of Figure **19** illustrates an example of Cd109 charge spectra in Time-Over-Threshold (TOT) units from four individual pixel channels. In the analysis of these data, fits were made to the photo peaks and using the TOT calibration information, the magnitude of the FE-I1 injection capacitors was determined. The

upper-right histogram shoes the distribution of all fitted photo peaks (in TOT units) while in the lower right plot the absolute input charge distribution is reconstructed

In Figure 20 an Am241 hitmap for an MCM is shown in which four FE receivers channels were enabled on the MCC. The FEs in the case were tuned to a common threshold of 2500e-. Along with verifying the absolute calibration of the injection capacitors which is imperative to understanding the analogue front-end properties, tests of this kind are useful in determining bump bond yields. It is also important to verify that the module is well behaved (in terms of e.g. noise occupancy) when all of the pixels are switched on in unison. Measurements of this kind have been used to show that fully instrumented MCMs have a mean noise occupancy which is less that 10^{-8} trigger^{-1} pixel^{1} once the thresholds of all pixels have been carefully tuned in order to optimise the threshold matching.

5.4 COMPONENT TRACKING

The Pixel production data, including bare module components information, are being stored in the Oracle 8i database (PDB). All the components have ATLAS unique serial numbers[1] as the PDB primary keys. The data structure used in the PDB corresponds to the Pixel detector mechanical assembly structure. Figure 21 represents an example of bare module prototype (ATLAS serial number 20210210111302) assembly, composed of sensor tile and 16 FE chips, with links to more detailed information about components and tests. The order of FE chips (N) in the bare module with respect to the sensor tile HV pad is shown in Figure 22.

Module number	Module type	Test list	Item number	N	Item type	Test list	Creator	Assm Date
-								
20210210111302	BareFEIModuleProt	Tests: Gen.info	20210110111302	1	SensorTile	Test: Gen.info	P_BON	17-JUL-02
			20211000809031	0	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211000810101	1	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001009111	2	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001008081	3	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001011081	4	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001002091	5	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001011041	6	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001002051	7	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001002041	8	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001005111	9	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001007091	10	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001001051	11	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001001061	12	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211001002101	13	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211000811091	14	FEchip	Test: Gen.info	P_BON	17-JUL-02
			20211000810031	15	FEchip	Test: Gen.info	P_BON	17-JUL-02

Figure 21: *Database summary table example.*

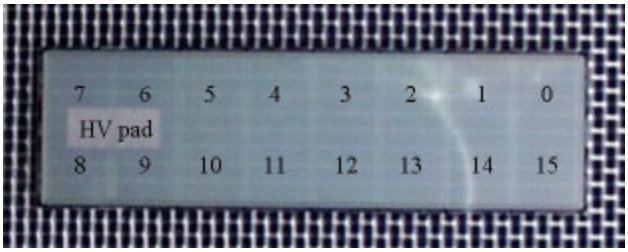

Figure 22: *numbering of FE chips on a bare module looking from top (sensor side).*

The item table (see example in the next figure) contains bare module and components information about the quality (passed: yes/no), location and user to which this item belongs, the current state (assembled:yes/no), manufacturer, the name given by manufacturer, amount available (for equal items) and the date of registration. The information is being uploaded via a specific java application, which performs different cross-checks to reduce the probability of operator error.

Serial number	Mfr Ser number	Item Type	Stocks	Date	Manufacturer	Passed	Location name	Assembled	Owner
-									
20210210111302	4455-13	BareFEIModuleProt	1	17-JUL-02	IZM	YES	University of Bonn	YES	P_BON

Figure 23: *Item table.*

To facilitate the part tracking the serial numbers carry some useful information about their origin. The FE chip and sensor tile serial numbers point to the wafer serial number. The bare module serial number points to the serial number of the sensor tile[2].

The PDB report facility allows selections of parts and tests by different parameters like type, date, laboratory name, manufacturer name, etc, which helps to keep the production process under control. The PDB is being updated permanently with the shipment data as well.

The FE chip serial number is derived from its position on the wafer. Once the wafer is cut, however, there is no unique marking on each chip. The way chips are tracked is by returning them to bump vendors after dicing and reprobing in labelled carriers (such as gel packs) which have a polarized orientation. We require that the vendors document the location of each chip flipped onto each module. Unfortunately there is no failsafe method for verifying that the information reported to us by the vendors is correct, but we have concluded that the consequences of potential occasional mistakes in chip tracking are not serious enough to require a failsafe procedure.

6 PRODUCTION ORDER AND TENDER DOCUMENT

The bare module production will start with a tender to the bumping vendors around January of 2003. Estimated time for the whole tender process until the production can start is 6 months, so the full production can start in July of 2003.

The actual "Tender" defining the Technical Specification of ATLAS Pixel Bare Modules and the contractual issues of the production order is a self-contained separate document. The scope of this document is not only to define the special technical requirements of the Bare Modules, like bump diameter, pitch and defect rate, but also to define the quantity and delivery items of the order. Also the quality control of the delivered modules is an important part of the Tender Document. Below is a short description of the key items covered in the document.

6.1 QUANTITY OF THE ORDER

The total number of bare modules in the completed detector is 1154 for a 2-hit system and 1744 for 3-hit system. Taking required spares and yield losses into account the bare module numbers to be purchases are shown in Table 4. It is included in this document for illustration only. The tender document itself is the official source of the order quantity. Similarly the breakdown of the spare requirements and yields at each step is maintained as a separate document, ATLIP….

Category	2-hit Detector	3-hit Detector
In the detector	1154	1744
Functional Spares	173	236
Yield Loss	216	324
Total to be Purchased	1531	2304

Table 4: *Number of bare modules needed. The numbers or yield and functional spared have compounded all assembly steps that use bare modules. For example spare sectors and staves as well as individual spare modules. This table is for illustration only and the numbers given may differ somewhat for the actual quantities eventually ordered and used.*

Assuming 50% good FE die per wafer and 70% good tiles per sensor wafer the number of wafers to be processed by the bump vendors is thus 165 (246) IC and 735 (1098) sensor wafer for a 2-hit (3-hit) system. Additionally, the vendors must be able to comply with an increase of the final production volume by up to a factor 1.5 within the contract.

6.2 DELIVERY SCHEDULE OF THE ORDER

pre-production	Jan 2003 – June 2003
start of production	July 2003
planned duration of delivery	18 months

The firms must accept a possible slippage of the planned schedule by as much as 1 year at no additional cost.

6.3 SPECIAL TECHNICAL REQUIREMENTS

Besides the requirements coming from the physics performance of the ATLAS Pixel detector like the total number of bump defects, there are some requirements which deal with the further production of pixel modules and special materials like the use of high resistivity silicon. The requirements mentioned in the tender document are:

- Maximum process temperature: 260°C
- It is required to explain the technology process steps to ATLAS pixel group
- Further processing on the module assembly (wire bonding, gluing and component mounting) will be necessary. The company must guarantee that the various process steps do not impose any serious obstacles that prevent further treatment using standard tools, remnants of photo resist or other deposits, etc. In such cases the necessary cleaning steps must be provided at no additional costs.
- Recording and book keeping of dies and wafers
- Firms must specify all processing temperature cycles

6.4 QUALITY CONTROL

Quality is the most important issue to be specified in the contract with the vendors. Therefore detailed specification of test procedure and acceptance criteria is included in the Tender Document. The quality control will be done by various institutes within the ATLAS Pixel collaboration. It consists of X-ray inspection, which can be provided by the vendor, and electrical measurements.

The test steps applied to bare module and the acceptance criteria are as follows:

1. After reception of the module (bare module) X-ray inspection is performed or the inspection maps provided by the vendor (if applied) are checked. Modules with more than 50 individual bump failures are rejected.
2. The electrical performance of module ICs will be tested with dedicated equipment and minimal human intervention. An IC is considered "bad" if it fails the standard ATLAS test procedure applied also to IC wafers and ICs after thinning and dicing. No modules with "bad" ICs are accepted.

3. Accepted modules will be wire bonded to flex kapton hybrids (FLEX) using standardized wire bonding equipment. Modules must be bondable using standard bonding parameters.

4. The module is then fully assembled using dedicated equipment (module assembly) in order to allow further electrical testing. The electrical measurement must confirm the absence of electrical connection detailed as:
    a. the proper operation of each electronics channel will be verified by charge injection
    b. the input capacitance of each channel will be determined using the noise/capacitance correlation.
    c. a gamma radioactive source will irradiate the whole detector and all the channels will be read-out.

    These tests can identify disconnected bumps, merged bumps, and other problems such as residue between bumps which cal lead to degraded performance. Modules with more than 150 total bump defects, or more than 30 in a single chip, will be rejected.

---

[1] *ATC-OQ-QA-2040, ATLAS Part Identification*

[2] *ATL-IP-MN-0002, Part Numbering in the ATLAS Pixel Production Database*